

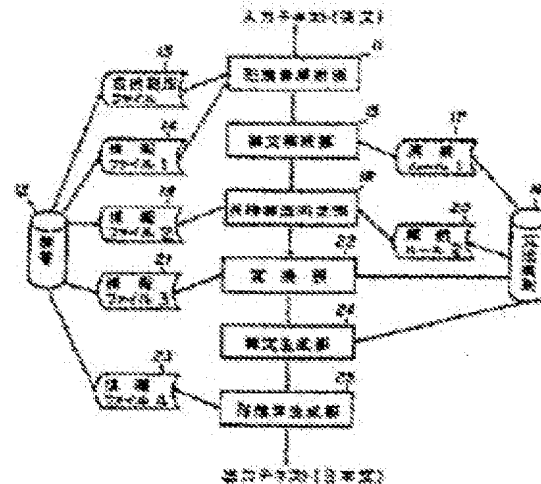
MACHINE TRANSLATION DEVICE

Publication number: JP4107769
 Publication date: 1992-04-09
 Inventor: OGURO YOSHIHISA
 Applicant: RICOH KK
 Classification:
 - international: G06F17/28; G06F17/28; (IPC1-7): G06F15/38
 - European:
 Application number: JP19900227748 19900829
 Priority number(s): JP19900227748 19900829

Report a data error here

Abstract of JP4107769

PURPOSE: To generate translated sentence, which is faithful to the conceptional structure of a source sentence and can be easily comprehended, by detecting indeclinable words, declinable words and case postpositional particles in the clause of the translated sentence and changing the order of words in the translated sentence according to the detected results. **CONSTITUTION:** This device is equipped with a morpheme analysis part 11, dictionary 12, inflective word form file 13, information file 14, syntax analysis part 15, grammatical rule 16, analysis rule 17, English construction identifying part 18, information file 19, analysis rule 20, information file 21, conversion part 22, information file 23, syntax generation part 24 and morpheme generation part 25. The source sentence of an objective language is translated into an objective language by using the analysis dictionary 12, and the inputted source sentence and the translated result are displayed. Then, declinable parts, declinable parts and case postpositional particles (corresponding phrases) in the clause of the translated sentence are detected and according to these detected results, the order of words in the translated sentence is changed. Thus, the natural and easily comprehensible translated sentence is generated without losing the meaning of the conceptional structure of the original.



Data supplied from the esp@cenet database - Worldwide

⑨ 日本国特許庁(JP)

⑩ 特許出願公開

⑧ 公開特許公報(A) 平4-107769

⑪ Int. Cl.³

識別記号

庁内整理番号

⑫ 公開 平成4年(1992)4月9日

G 06 F 15/38

T

9194-5L

審査請求 未請求 請求項の数 3 (全6頁)

⑬ 発明の名称 機械翻訳装置

⑭ 特 願 平2-227748

⑮ 出 願 平2(1990)8月29日

⑯ 発 明 者 大 黒 慶 久 東京都大田区中馬込1丁目3番6号 株式会社リコー内
⑰ 出 願 人 株 式 会 社 リ コ ー 東京都大田区中馬込1丁目3番6号
⑱ 代 理 人 弁 理 士 高 野 明 近 外1名

明 細 書

1. 発明の名称

機械翻訳装置

2. 特許請求の範囲

1. 少なくとも、対象言語の原文を入力するための入力手段と、該入力手段により入力された原文を記憶する記憶手段と、言語解析に必要な知識を記憶している解析辞書と、前記記憶手段に記憶された対象言語の原文を前記解析辞書を用いて目的言語に翻訳する翻訳手段と、前記入力手段により入力された原文と前記翻訳手段により翻訳された翻訳結果を表示するための表示手段と、原文の文節内の体言及び用言、格助詞を検出する検出手段と、該検出手段の検出結果によって原文の語順を変更する語順変更手段とを有することを特徴とする機械翻訳装置。

2. 前記語順変更手段において、語順変更対象文節間に句読点を付与する付与手段を有することを特徴とする請求項1記載の機械翻訳装置。

3. 前記語順変更手段において、前記検出手段

の検出結果だけでなく、原文側の語順も反映した語順変更手段を有することを特徴とする請求項1記載の機械翻訳装置。

3. 発明の詳細な説明

装置分類

本発明は、機械翻訳装置に関し、より詳細には語順整形機能を有する機械翻訳装置に関する。

従来技術

従来、ある言語を生成するための手段としては、ある概念を表わす抽象構造（概念構造）から、概念構造中の品詞（名詞、動詞、形容詞、副詞など）間の関係を基に表層の単語列を決定していた。しかしながら自然な文章を生成するには、いわゆる文法上の品詞分けでは不十分であり、当該言語において適切な語順整形を使用者が行う必要があった。

また上記の生成方法では、修飾関係があいまいになる場合があり、原文の概念構造の意味を忠実に表現する訳文（表層の単語列）を必ずしも生成できなかった。

目 的

本発明は、上述のごとき実情に鑑みてなされたもので、原文の概念構造の持つ意味を失わないまま、しかも自然で理解容易な訳文を生成するために、語順変形機能を有する機械翻訳装置を提供することを目的としてなされたものである。

構 成

本発明は、上記目的を達成するために、(1)少なくとも、対象言語の原文を入力するための入力手段と、該入力手段により入力された原文を記憶する記憶手段と、言語解析に必要な知識を記憶している解析辞書と、前記記憶手段に記憶された対象言語の原文を前記解析辞書を用いて目的言語に翻訳する翻訳手段と、前記入力手段により入力された原文と前記翻訳手段により翻訳された翻訳結果を表示するための表示手段と、訳文の文節内の体言及び用言、格助詞(相当句)を抽出する抽出手段と、該抽出手段の抽出結果によって訳文の語順を変更する語順変更手段とを有すること、更には、(2)前記語順変更手段において、語順変

更対象文節間に句読点を付与する付与手段を有すること、更には、(3)前記語順変更手段において、前記抽出手段の抽出結果だけでなく、原文例の語順も反映した語順変更手段を有することを特徴としたものである。すなわち、本発明では、原文の文節内の体言及び用言、格助詞(相当句)を抽出し、その抽出結果によって訳文の語順を変更する語順変更手段と、依存関係のない修飾関係になり得る文節間に読点を付与する付与手段とを有することを特徴としたものである。これによって、あいまいな修飾関係を排除することができ、原文の概念構造に忠実で、かつ理解容易な訳文を生成することができる。以下、本発明の実施例に基づいて説明する。

第1図は、本発明による機械翻訳装置の一実施例を説明するための構成図で、ここでは英日機械翻訳装置の組合について説明する。以下、特にこととされない限り、英日機械翻訳装置に限定するが、本発明による方式は語順の制約が少ない文法体系を持つ言語(母語系)に対して一般的に成

立することである。図中、1.1は形態素解析部、1.2は辞書、1.3は国語語彙ファイル、1.4は情報ファイル1、1.5は構文解析部、1.6は文法規則、1.7は解析ルール1、1.8は英語構造限定部、1.9は情報ファイル2、2.0は解析ルール2、2.1は情報ファイル3、2.2は変換部、2.3は情報ファイル4、2.4は構文生成部、2.5は形態素生成部である。機械翻訳システム装置中の辞書1.3には、語形、活用形、品詞、訳語などが記述されている。

入力された英文は、まず形態素解析部1.1で一つ一つの単語として認定され、番号+姓や数字+単位などの特別な文字列は句としてまとめられる。

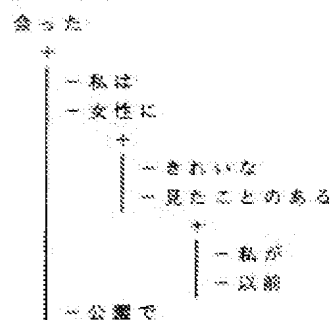
構文解析部1.5で入力英文の構造が解析される。文末から始めて右から左に、辞書記述内容と文法ルールを適用することにより文構造を決定していく。例えば、前置詞+冠詞+名詞が前置句と解析される。

英語構造限定部1.8では、係り受け関係等を明らかにするために動詞ノードを源とする木構造が

作られ、前置句の係り先等が決まる。次に、主起動詞の成分が調べられ、適切な訳語が取られる。

前記木構造は変換部2.2で日本語の言い回しを組み合わせ、中間構造に変換される。構文生成部2.4では文節単位で日本語を生成した後、形態素生成部2.5で助詞の並びを決定したり、用言の活用形を替える。

変換部2.2からの出力は、以下のような依存関係を表現している。



このような依存関係から、訳語文を生成する。日本語は、文節内の語順は厳密に決まっているが、文節の順に関しては比較的自由である(述語が文

尾、修飾文節は被修飾文節に先行する。など）、したがって以下の文が生成されたとする。

①私は、私が以前見たことのあるきれいな女性に公園で会った。

②私は、きれいな私が以前見たことがある女性に公園で会った。

③私は、公園で私が以前見たことのあるきれいな女性に会った。

生成された表層文①の係属関係は依存関係を忠実に表現している。一方、表層文②③においては、複数の係受け関係が可能であり、誤った依存関係に解釈されやすい。例えば②においては、「きれいな」のは「私」と解釈されやすく、③では「公園で」「以前見たことのある」「女性」に「会った」と解釈されやすい。この理由は、修飾文節は、修飾可能な文節の中で最も近くの文節を修飾する傾向があるからである。

原文の概を構造的持つ意味を失わないまま、しかも自然で理解容易な訳文を生成するためには、訳文内のあいまいな係属関係を排除する語順整理

機能が必要である。日本語では、長々複雑な文節は他の文節に先行させ、文節の位置に配置したほうがよい。これは、主題や他の文節と語順の間に構造的複雑な文節が置かれると、複雑な句に先行する文節の構文上の係り関係がわかりにくくなり、かつ読み難くなるからである。

構造的複雑な文節と判定する方法は、当該文節内の修飾語（相当句）を移えればよい。また連体修飾の場合、修飾文節と被修飾文節との間に他の体言があると連体修飾文節がその体言に係りやすくなり、誤った依存関係に解釈してしまう。同様に連用修飾の場合でも、修飾文節と被修飾文節との間に他の用言があると連用修飾文節がその用言に係りやすくなる。

以上の点を考慮して、具体的な実現手段を示す。次に示す原則にしたがって語順を調整する。いずれも兄弟関係の語順に関する原則である。ここでいう兄弟関係とは、依存関係において、同じ文節を修飾する、同レベルの文節の関係を指す。

原則としては可能な限り修飾文節は被修飾文節

(2.2.5) 体言+の

フランスの、車の、男性用の

連体修飾関係における語順規則は以下の通りである。

(1) 体言密接グループ

それが係る体言の近くに置く。体言密接グループ内での語順は、文節長が短いものほど体言に近い。このグループ間には読点を付与しない。文節の長さとしては、文字数や音節数などが考えられる。

(2) 体言非密接グループ

上記「体言密接グループ」内の語順を決定した後、その体言を含んだ文節を修飾する。

修飾される体言を含んだ文節から遠い順に

(2.2.a) (2.2.b) (2.1.b) (2.1.a) → (体言を含む文節)
体言を含む 体言を含まない

の順に並べる。同グループ内では文節長の短いものが体言文節に近い。

あいまいな係り受け関係を避けるために、隣接する文節間に以下の読点規則を設ける。

の直前が浸ましい、小さい範囲（内側）の係り受けから広い範囲の文節の係り受けへと、順に決定していく。接続詞をまたがって語順を変えることはない。

まず、連体修飾関係について説明する。連体修飾語（節）となる文節を次のように分類する。

連体修飾語となる文節の分類

(1) 体言密接グループ

形容詞（連体形）→ 赤い、大きい

連体詞 → この、あらゆる、大きな

形容動詞（連体形）→ 暖かな、静かな、→ 可能な

(2) 体言非密接グループ

(2.1) 体言を含まない

(2.1.a) 動詞（連体形）

→ 許せる、走る、獲られた

(2.1.b) 連用修飾語+動詞（連体形）

→ 速く走る、優しく接れる

(2.2) 体言を含む

(2.2.a) 修飾語を含むもの

→ 効率が良い、息を切らした

① 読点を付与しない

「体言を含む文節」「体言を含まない文節」

例 効率がよい速く走る→車

「体言を含まない文節」「体言を含む文節」

例 速く走る車買られた→車

② 読点を付与する

「体言を含む文節」「体言を含む文節」

例 効率がよい フランスの→車

「体言を含まない文節」「体言を含む文節」

例 速く走る フランスの→車

次に、連用修飾関係について説明する。連用修飾語(節)となる文節を次のように分類する。

連用修飾語となる文節の分類

(1) 用言密接グループ

副詞 … ぶと、さらに、いくぶん

形容詞 (連用形) … 赤く、大きく、速く

形容動詞(連用形) … 暑かに、寒さに

(2) 用言非密接グループ

(2.1) 用言を含まない

体言+格助詞(相当句)→インクで、東京駅から

(2.2) 用言を含む

(2.2.a) 格要素を含むもの

… 行儀がよく、息を切らして、

(2.2.b) 動詞(連用形)+後続助詞

… 眠ずかしがって、食べながら、

(2.2.c) 用言を含む名詞句+格助詞(相当句)

… 赤いインクで、速く走る車によって

連用修飾関係における語順原則は以下の通りである。

(1) 用言密接グループ

それが係る用言の近くに置く。用言密接グループ内での語順は、「程度を添わず節親」が、それが修飾する用言に最も近く、その他の文節長が短いものほどそれが修飾する用言に近い。このグループ間には読点を付与しない。

(2) 用言非密接グループ

「用言密接グループ」内の語順を決定した後、その用言を含んだ文節を修飾する。

修飾される用言を含んだ文節から遠い順に

(2.2.a) (2.2.b.c) (2.1) → (用言を含む文節)
用言を含む 用言を含まない

の順に並べる。同グループ内では文節長が短いものほど、用言文節に近い。

あいまいな係り受け関係を避けるために、隣接する文節間に以下の読点原則を設ける。

① 読点を付与しない

「用言を含む文節」「用言を含まない文節」

例 行儀がよく箸で→食べる

「用言を含まない文節」「用言を含まない文節」

例 東京駅から車で→来る

② 読点を付与する。

「用言を含む文節」「用言を含む文節」

例 行儀がよく、敬語しながら→食べる

「用言を含まない文節」「用言を含む文節」

例 電話によって、食べながら→連絡する

先に示した依存関係に本発明の語順変更手段を施した例を示す。最初に、最も下位の依存関係の語順が決められる。その結果、

「私が」「以前」→「見たことのある」

という語順になる。

次に、文節「女性に」への連用修飾語は、「きれいな」「私が以前見たことのある」の二つである。「きれいな」は連体形で体言密接グループ、「私が以前見たことのある」は格要素を含むので、体言非密接グループである。よって語順は、「私が以前見たことのある」「きれいな」→「女性に」となる。

最後に、文節「会った」への連用修飾語は、「私は」「私が以前見たことのあるきれいな女性に」「公園で」の三つである(主語も連用修飾語の一種と考える)。

「私は」と「公園で」は、用言非密接グループの(2.1)である。「私が以前見たことのあるきれいな女性に」は、用言非密接グループの(2.2.a)である。したがって語順は、

「私が以前見たことのあるきれいな女性に」「公園で」「私は」→「会った、」

となる。

最終的に生成される表層文は、

「私が以前見たことのあるきれいな女性に公園で
私は会った。」
となる。

第2図は、以上に説明したような語順整形手順
のフローチャートである。以下、各ステップに従
って順に説明する。

step 1: まず、処理対象を移動する。

step 2: 当該文節の文節グループ判定を行う。

step 3: 兄弟関係の文節があるかどうか判断する。

step 4: 語順規則に従って決定する。

step 5: 読点規則に従って読点付与を行う。

step 6: 親子関係があるかどうか判断する。親子
関係がなければ終了し、あればstep 1に戻る。

本発明による文節の分類手段、すなわち文節中
の体言・用言の存在の判定は容易に実現できる。
まず第1表のように辞書中の語語情報に体言・用
言の存在を記入しておく。また格要素の有無も同
時に付与しておく。

文節生成時には、これらの情報を受け継いでい
くことにより、生成された文節に対しても判定す

"He always calls and scolds his son."

「彼は彼の息子をいつも呼んで、そして叱る。」
助動詞の並列の場合も同様である。

"He can and even do it well."

「かれはそれを上手にでき、そしてしなければな
らない。」

前記の「でき」のように連用形に活用せず、終
止形のまゝにする方法もある。

「彼はそれを上手にできる。そしてしなければな
らない。」

英文例での特殊な出現位置

英文例において文節部にあり、和文例で連用修
飾となるものは、和文例でも先頭に読点を挿入
て訳出する。

"He was absent from school yesterday."

「彼は学校を昨日休んだ。」

"Yesterday he was absent from school."

「昨日、彼は学校を休んだ。」

また、英文例で副詞の両側をコンマで区切られ
ていたら、和文例では読点を挿入してその文の先頭

ることができる。

第1表 評語記述例

見出し語	訳語	格助詞数	体言	用言	音節数
efficient	有効な	0	無	有	5
	効果がある	1	有	有	6

次に原文例の情報を反映した語順変更手段につ
いて説明する。

並列（等位）構造

並列構造の場合、並列構造の範囲内の語順は原
文に基づき、読点は付与しない。例えば、

"beautiful brother-in-law and John"

→「美しい（義理の兄弟とジョン）」

→「（美しい義理の兄弟）とジョン」

の場合、英文例でも複数の修飾関係になり得るの
で、原文のあいまい性を訳文においても保存する。
したがって上記各訳句の訳は

「美しい義理の兄弟とジョン」

になる。

共通補語を持つ並列の場合も同様に、語順は原
文に従う。

に訳出する。

"Tough he studied hard, he (unfortunately)
failed in his exam."

「彼は一所懸命勉強したけれども、不幸にも、彼
は彼の試験に落ちた。」

複合名詞

英語例でも同連綴でありながら、和文例では、
連体修飾のために「の」を挿入している場合には、
いかなる場合でも名詞連綴の範囲内に修飾成分は
挿入しない。

慣用的な表現

慣用的表現では述語の直前に連用修飾成分を置
くと奇異な日本語になる場合があるので、このよ
うな場合には、慣用的な表現の範囲の直前に修飾
成分を配置する。

"He sometimes takes fun of old man."

「彼は老人をときどき馬鹿にする。」

「馬鹿にときどきする」とは言わない。

本発明は以上の実施例に限定されるものではな
く、訳文の文節内の体言及び用言、格助詞（相当

10) を検出し、その検出結果によって原文の語順を変更する語順変更手段と、依存関係にない修飾関係になり得る文節間に読点を付与する付与手段とを有し、これによって、あいまいな修飾関係を排除し、原文の類似構造に忠実で、かつ理解容易な訳文を生成することが可能な機械翻訳装置をもつ機械翻訳システムを提供することにある。

以上の説明から明らかなように、本発明によると、以下のような効果がある。

(1) 請求項1、2に対応する効果：原文におけるあいまいな修飾関係を排除し、原文の類似構造に忠実で、かつ理解容易な訳文を生成することができる。

(2) 請求項3に対応する効果：原文側の語順のある程度反映しているため、原文と訳文との間での印象があまり変化せず、したがって読者間で誤合しやすくなる。

4. 図面の簡単な説明

第1図は、本発明による機械翻訳装置の一実施

例を説明するための構成図。第2図は、語順整形手段のフローチャートである。

11…形態素解析部、12…辞書、13…屈折形態ファイル、14…情報ファイル1、15…構文解析部、16…文法規則、17…解析ルール1、18…英語構造部定部、19…情報ファイル2、20…解析ルール2、21…情報ファイル3、22…変換部、23…情報ファイル4、24…構文生成部、25…形態素生成部。

特許出願人 株式会社 リコー
代理人 高野 明 近
(ほか1名)

図1 100

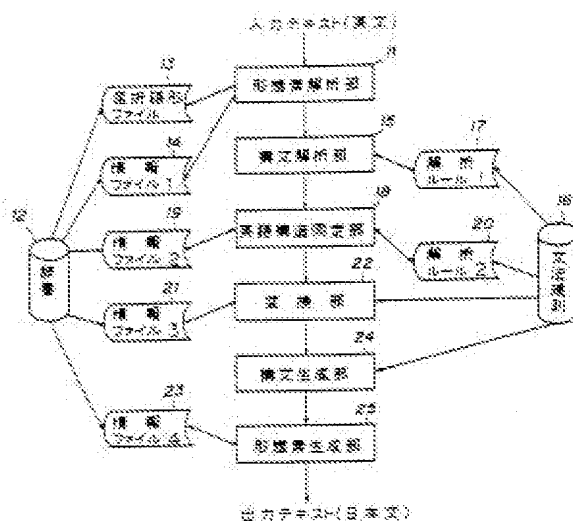


図2 100

